Ogden Air Logistics Center

U.S. AIR FORCE

Trust on the Battlefield in an Age of Automation

Daniel P. Stormont
Embedded Systems Engineer
520 SMXS/MXDEA
(801) 775-3191
daniel.stormont@hill.af.mil

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
APR 2010		**00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Trust on the Battlefield in an Age of Automation	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Ogden Air Logistics Center,520 SMXS/MXDEA,Hill AFB,UT,84056**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the 22nd Systems and Software Technology Conference (SSTC), 26-29 April 2010, Salt Lake City, UT.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**35**	

Overview

- Why Trust Matters
- Autonomy Defined
- Trust Defined
- Identifying Elements of Trust
- Simulating Trust
- Human-Robot Trust Games
- Web-based Game
- Diplomacy Agent
- Identifying Human-Agent Trust KPPs

Why Trust Matters

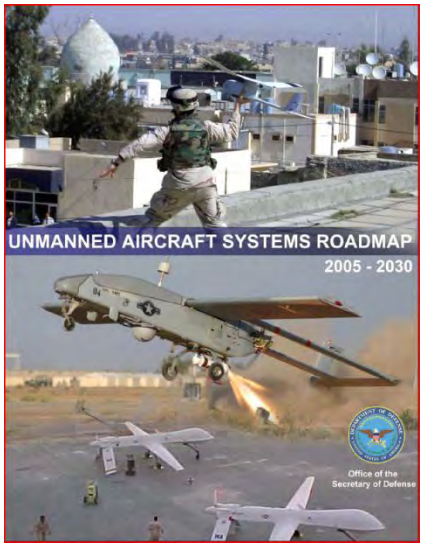
- Congressional Mandate
  - The National Defense Authorization Act for Fiscal Year 2001, Public Law 106-398, Congress mandated in Section 220 that “It shall be a goal of the Armed Forces to achieve the fielding of unmanned, remotely controlled technology such that… by 2015, one-third of the operational ground combat vehicles are unmanned.”[1]

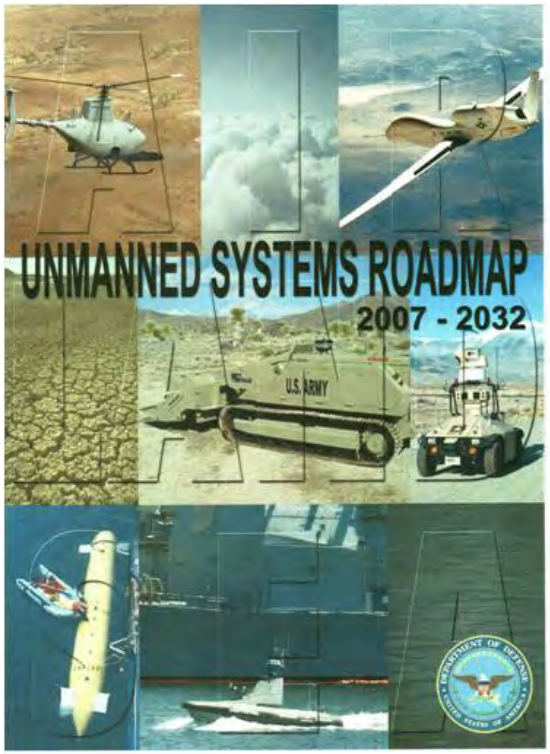
Why Trust Matters

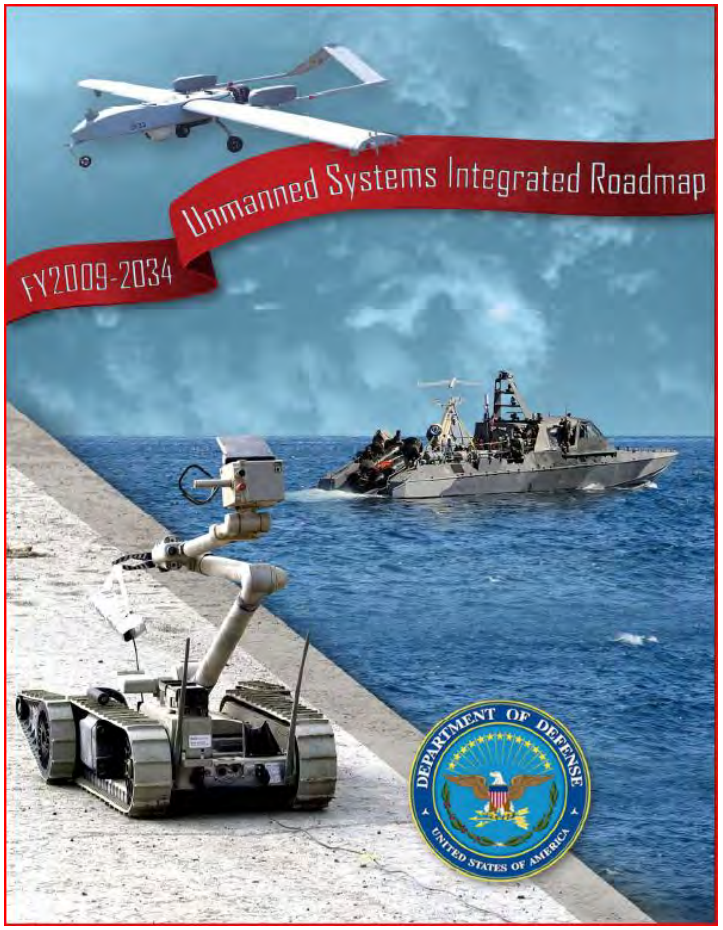
- Secretary of Defense Gates' Comments
  - On 21 April 2008, Secretary Gates made the following comment about unmanned systems

"Unmanned systems cost much less and offer greater loiter times than their manned counterparts, making them ideal for many of today's tasks. Today, we now have more than 5,000 UAVs, a 25-fold increase since 2001. But in my view, we can do and we should do more to meet the needs of men and women fighting in the current conflicts while their outcome may still be in doubt. My concern is that our services are still not moving aggressively in wartime to provide resources needed now on the battlefield. I've been wrestling for months to get more intelligence, surveillance and reconnaissance assets into the theater. Because people were stuck in old ways of doing business, it's been like pulling teeth."[2]

Why Trust Matters







3

Why Trust Matters

- President's Budget for Unmanned Systems[3]

PORs FY09PB ($M)	Funding Source	FY09	FY10	FY11	FY12	FY13	TOTAL
UGV	RDT&E*	$1291.2	$747.5	$136.2	$108.7	$68.9	**$2,353**
	PROC*	$33.4	$42.3	$53.5	$59.5	$21.1	**$210**
	O&M*	$2.9	$3.9	$3.0	$12.8	$10.1	**$33**
UAS	RDT&E	$1347.0	$1305.1	$1076.4	$894.0	$719.5	**$5,342**
	PROC	$1875.5	$2006.1	$1704.7	$1734.3	$1576.2	**$8,897**
	O&M	$154.3	$251.7	$249.0	$274.9	$320.2	**$1,250**
UMS	RDT&E	$57.3	$73.8	$63.2	$70.1	$76.9	**$341**
	PROC	$56.7	$78.4	$95.9	$91.6	$103.7	**$426**
	O&M	$5.0	$4.5	$11.3	$13.5	$13.9	**$48**
TOTAL		**$4,823**	**$4,513**	**$3,393**	**$3,260**	**$2,911**	**$18,900**

* RDT&E = Research, Development, Test, and Evaluation; PROC = Procurement; O&M = Operations and Maintenance

Why Trust Matters

- Common Performance Envelope Across Domains[3]

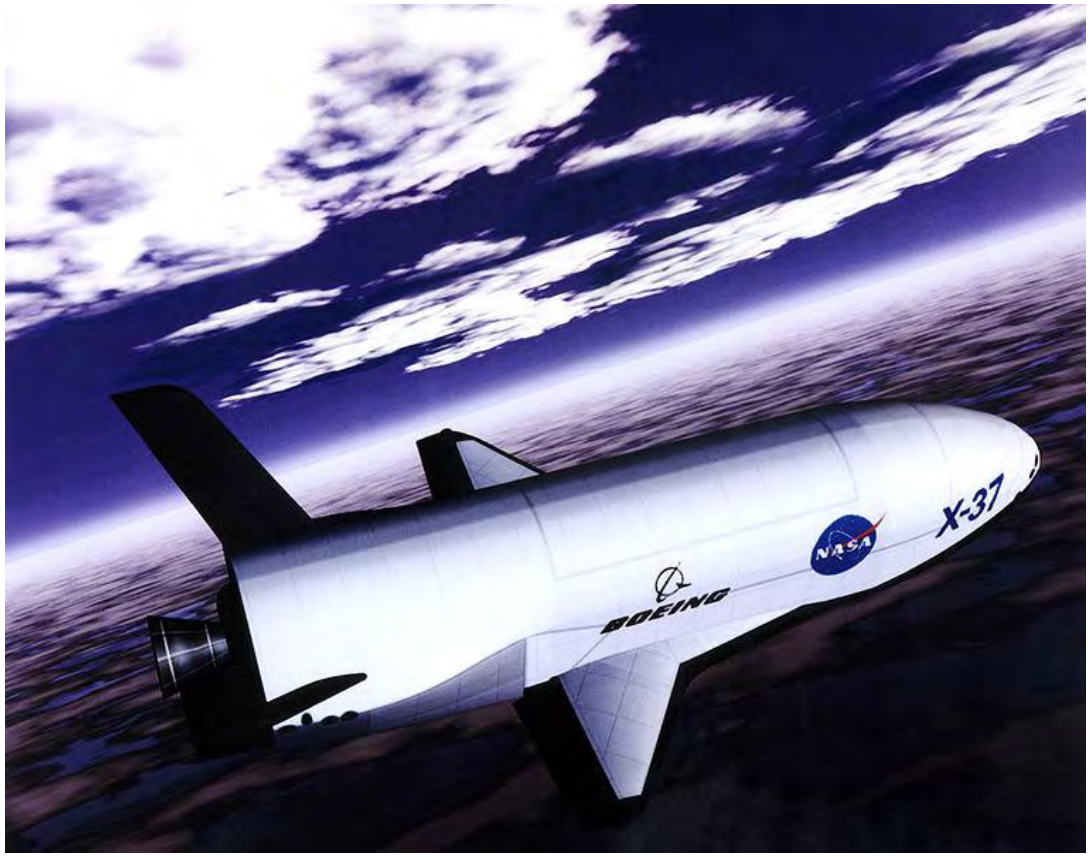
	2009 Evolutionary Adaptation	2015 Revolutionary Adaptation	2034
Commands	Physical Human Machine Interfaces	Scripted Voice Command/Hand Signals	Natural Language Understanding
Collaboration	Individual System	Teaming w/in Domain Collaboration Across Domains	Teamed Collaboration
Frequency	Constrained RF	Frequency Hopping	Multi-Frequency Communications
Mission Complexity	Operator Controlled		Autonomous Adaptive Tactical Behaviors
Environmental Capability	Limited Environmental Difficulty	Expanded Environmental Difficulty	All-Weather Environmental Difficulty
Product Line	Mission Package Product Line Dependent		Product Line Independent
OPSEC	Signature High		Signature Low
Operational Control	1 Operator / Platform	1 Operator / Domain	1 Operator / Team
Bandwidth	Limited	Advanced Bandwidth Management	Autonomous Bandwidth
Mission Endurance	Hours	Days Months	Years
Maintenance	Operator		Automated
Awareness	Sensor Data	Situational Awareness	Actionable Information

Why Trust Matters

- P.W. Singer of the Brookings Institution stated the following in an article published in 2009
  - "*So, despite what one article called „all the lip service paid to keeping a human in the loop,"the cold, hard, metallic reality is that autonomous armed robots are coming to war. They simply make too much sense to the people that matter.*" [4]

Why Trust Matters

- From recent news, the X-37B Orbital Test Vehicle was launched on 22 April 2010 and received a fair amount of media attention
  - Who is in control?
  - What are its capabilities?
  - How long will it remain in orbit and what is it doing?



NASA

Why Trust Matters

- An article in the February 2010 issue of *Popular Mechanics* asked "Can We Trust Robots?"[6]
  - The primary point of the article was that humans may be too trusting of robots because they anthropomorphize them
  - We humans think we understand a robot's "thought processes" and "motivations"
  - This could become worse as robots become ubiquitous and as we begin to rely upon them more

Autonomy Defined

- The American Heritage Dictionary defines *autonomy* as
  - *"Independence."* [7]
- Wikipedia defines an *autonomous agent* as
  - *"A system situated in, and part of, an environment, which senses that environment, and acts on it, over time, in pursuit of its own agenda. This agenda evolves from drives (or programmed goals). The agent acts to change the environment and influences what it senses at a later time."*
  - *"Non-biological examples include intelligent agents, autonomous robots, and various software agents, including artificial life agents, and many computer viruses."* [8]

Autonomy Defined

- There are varying degrees of autonomy
  - None – System is completely manually controlled
  - Partial – Some functions are automated
  - Sliding – The amount of autonomy is selectable
  - Full – The system operates entirely without human control
- With any autonomy the user gives up some level of control
- Wikipedia has a good definition of the abilities of a fully autonomous robot
  - *"A fully autonomous robot has the ability to*
    - *Gain information about the environment*
    - *Work for an extended period without human intervention*
    - *Move either all or part of itself throughout its operating environment without human assistance*
    - *Avoid situations that are harmful to people, property, or itself unless those are part of its design specifications."* [9]

Autonomy Defined

- Examples of Degrees of Autonomy

Increasing Autonomy

No Autonomy (Remote-control EOD robot)

US Air Force Photo

Partial Autonomy (ISS assembly robot)

NASA Photo

Sliding Autonomy (UAS)

US Air Force Photo

Full Autonomy (ALCM)

Boeing Photo

Trust Defined

- The "classic" definition of trust comes from Diego Gambetta
  - **"trust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both *before* he can monitor such action (or independently of his capacity ever to be able to monitor it) *and* in a context in which it affects *his own* action. When we say we trust someone or that someone is trustworthy, we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him."[10]**

Research Objective

- Assuming that autonomous agents are coming to hazardous environments, like disaster areas and combat zones, how can we trust them?
- This research project is attempting to identify the factors that contribute to trust in autonomous agents, in order to develop a set of key performance parameters (KPPs) for trusted agents
  - These KPPs can be used by intelligent agent developers to verify the trustworthiness of their agents and to convince users of their trustworthiness

Research Objective

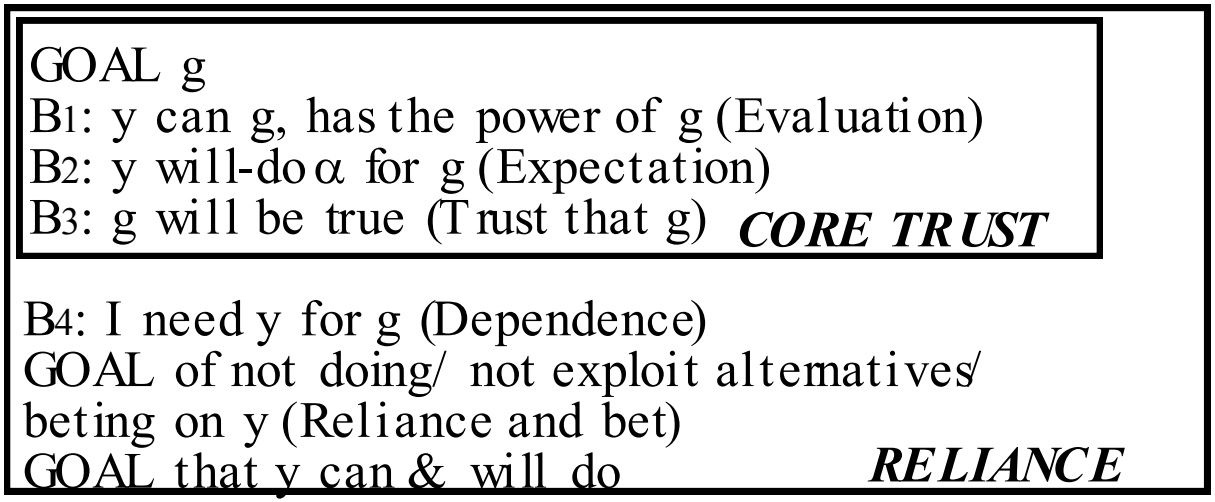
- Three elements to the research
  - Identify factors of trust published in the literature
  - Simulate human-agent interactions
  - Collect data on human interactions with autonomous agents
- The ultimate objective is to develop a process that can be validated in field trials

Identifying Trust Factors

- Cristiano Castelfranchi and Rino Falcone argue that

  Only agents endowed with goals and beliefs (cognitive agents) can "trust" another agent[11]

- They define the elements of trust with the following diagram



Mental ingredients of TRUST

Identifying Trust Factors

- Sarvapali Ramchurn, et al., identified two principle components of trust in an agent[12]
  - Confidence – Do I believe the agent can perform the desired task?
  - Reputation – Has this agent been successful in the past and have others trusted this agent, with good results?

Identifying Trust Factors

- Karen Fullam and Suzanne Barber focused on the importance of reputation when dealing with agents (either human or artificial) in the development of the ART Testbed[13]

Identifying Trust Factors

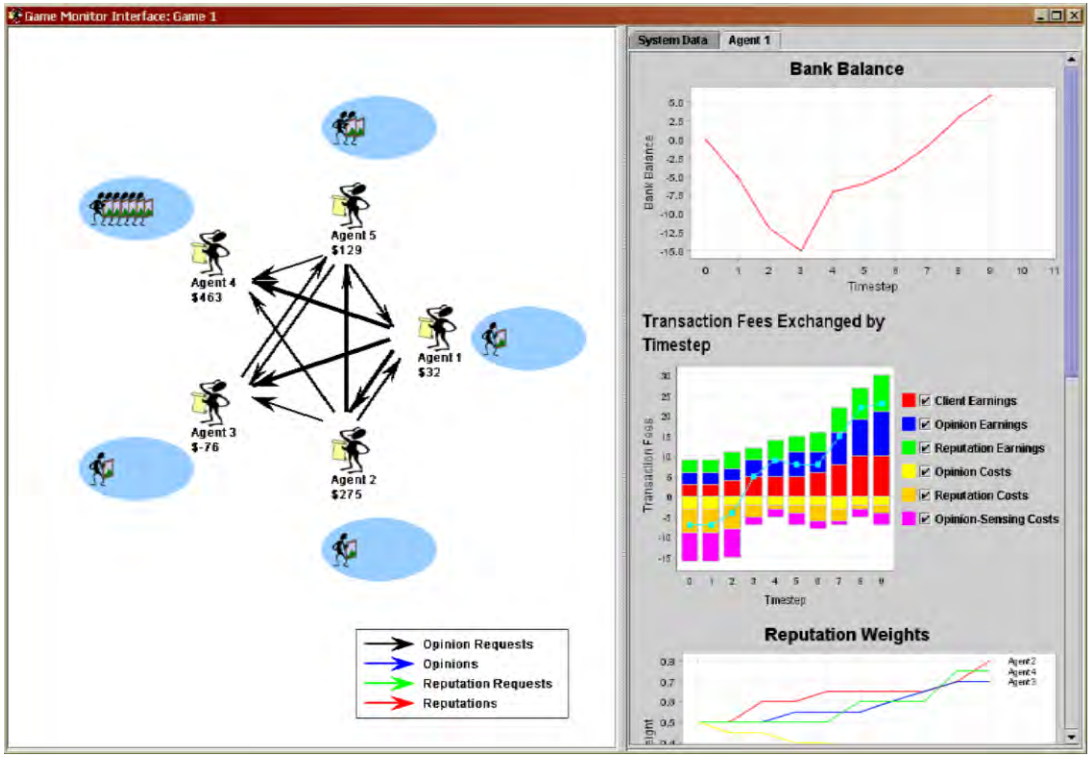
- Just in these three research papers, the following potential trust factors were identified
  - Evaluation
  - Expectation
  - Trust
  - Dependence
  - Reliance
  - Bet
  - Confidence
  - Reputation
- Which factors are the key factors?

Simulating Trust

- Trust of autonomous systems is an active area of research
- Much of the research uses simulations to explore issues of trust and to compare approaches
- The simulations tend to take two forms
  - Simulations designed specifically for examining issues of trust
  - Simulations that were designed for other areas of research that have been extended or adapted to trust research
- The following are some examples of both types of simulation

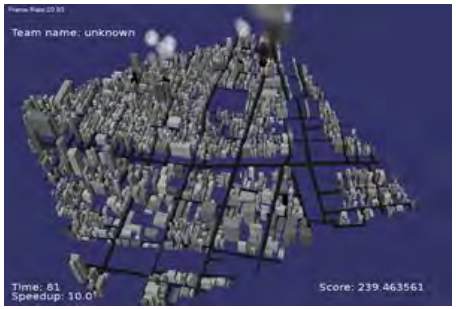
Simulating Trust

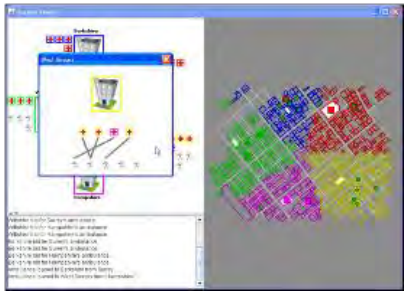
- Agent Reputation and Trust (ART) Testbed[14]
  - A simulation of multiple art appraising agents specifically designed for agent trust research



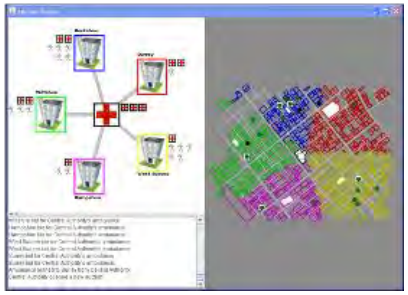
Simulating Trust

- RoboCup Rescue Simulation[15]
  - An agent simulation system that has been used for trust research[16]

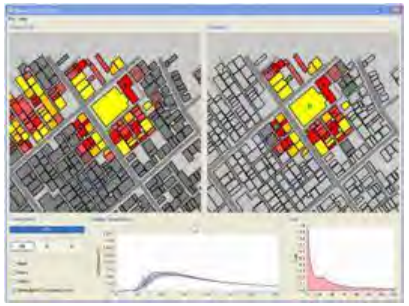




Coalition Formation



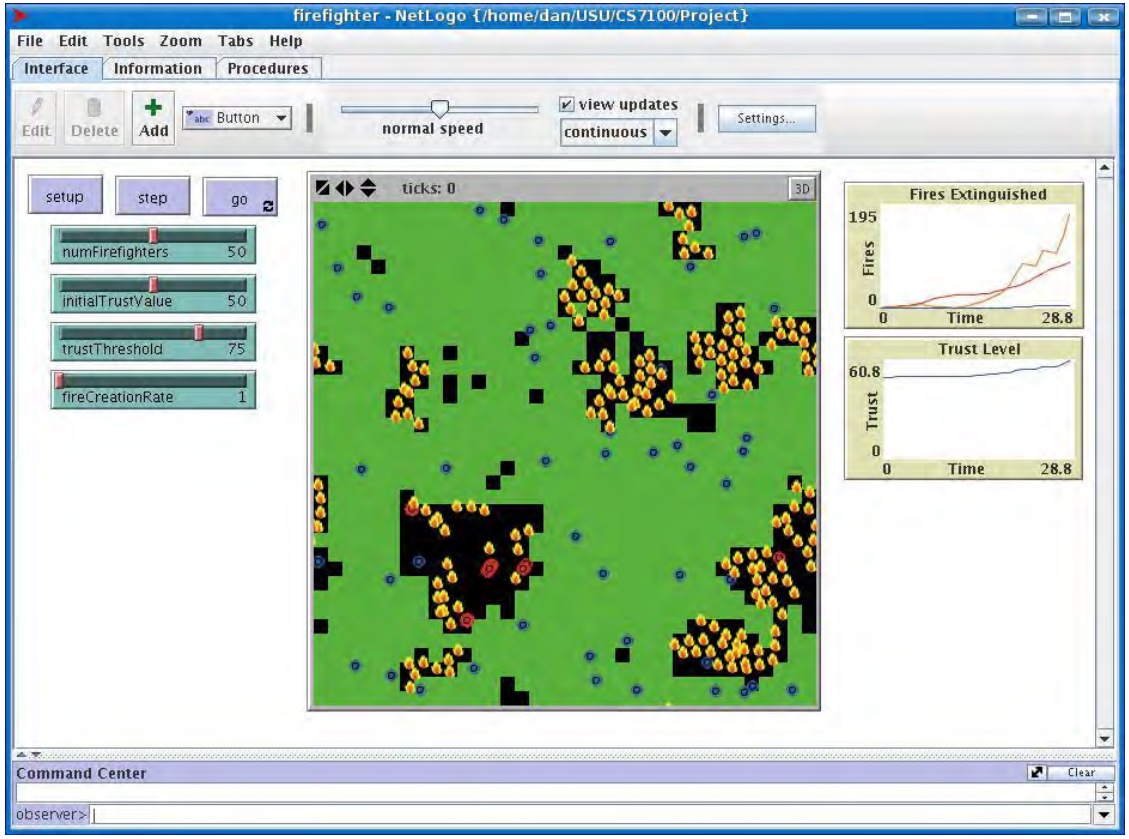
Resource Allocation



Estimation with Faulty Sensors

Simulating Trust

- Firefighting Simulation
  - A simulation of human-robot trust relationships



Human-Robot Trust Game Prototype

- A prototype of a game for collecting user interaction data was created in NetLogo
  - The objective of the prototype was to evaluate playability and design elements for data collection

Human-Robot Trust Game Prototype

- Last year a paper-based firefighting game was play tested at the RoboCup International Rescue Robotics Workshop
  - Based on feedback from the playtesting, a print and play game is being developed to gain wider feedback on the game mechanics
  - A print and play game simulating employment of UASs is also in preliminary development, with a working title of *Battle for Marjeh*
  - Links to both games and a wiki for feedback will be posted to the project web page when they are ready (see contacts slide at end of presentation)

Web-based Game

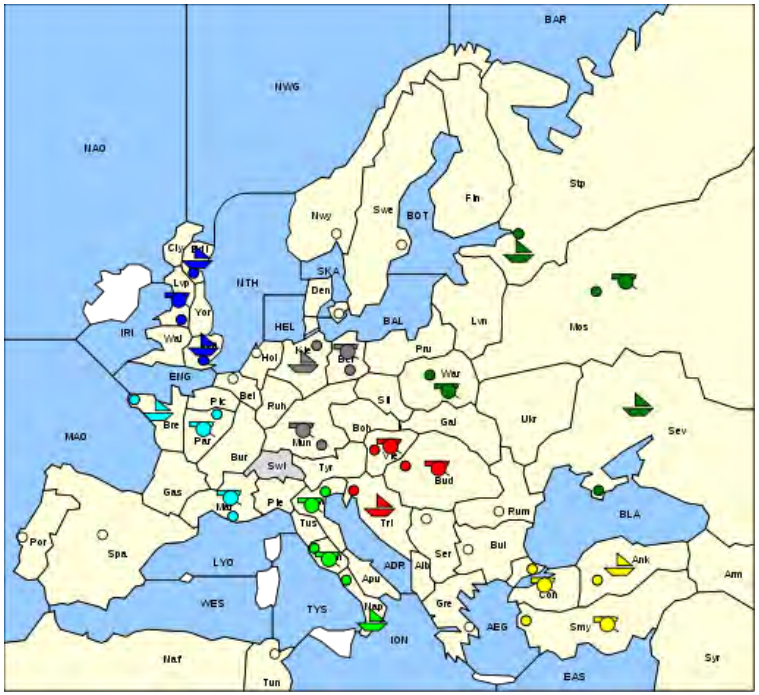
- To collect data from a variety of users, a web-based game will be developed, based on feedback from the print and play game, providing tools for collecting data on how game players utilize the robots

Diplomacy Agent

- As a parallel research effort, an autonomous *Diplomacy* playing agent is going to be developed to explore trust elements in an environment intended to encourage distrust
- *Diplomacy* games are available as play-by-email and via Internet play with a number of variants
- There is also a Diplomacy AI Centre in the UK that is open to all researchers to test *Diplomacy* agents in play against each other in a test environment called the Diplomacy AI Development Environment (DAIDE)[17]

Diplomacy Agent

- *Diplomacy* was created in 1954 by Allan B. Calhamer as a simulation of the diplomacy between the major powers prior to World War I in Europe
- The game typically has seven players
- In each turn, players participate in a negotiation phase, then all movements are resolved simultaneously[18]



Diplomacy Map by Martin Asal

Identifying Human-Agent Trust KPPs

- The final objective for this project is to develop Key Performance Parameters that can be used by autonomous agent designers
  - Serve as design guidelines for agents
  - Provide verification parameters for testing
  - Provide validation parameters for field testing

Summary

- Why Trust Matters
- Autonomy Defined
- Trust Defined
- Identifying Elements of Trust
- Simulating Trust
- Human-Robot Trust Games
- Web-based Game
- Diplomacy Agent
- Identifying Human-Agent Trust KPPs

Project Contact Info

daniel.stormont@aggiemail.usu.edu

Human-Robot Trust Project webpage:
https://sites.google.com/a/aggiemail.usu.edu/human-robot-trust-project/

Questions?

References

1. DARPA, *DARPA Urban Challenge.* Online: [http://www.darpa.mil/GRANDCHALLENGE/]
2. Robert Gates, Secretary Gates Remarks at Maxwell-Gunter Air Force Base, Montgomery Alabama, *DefenseLink News Transcript*, April 21, 2008. Online: [http://www.defenselink.mil/utility/printitem.aspx?print=http://www.defenselink.mil/transcripts/transcript.aspx?transcriptid=4214]
3. OSD, *Unmanned Systems Roadmap, 2009-2034,* April 6, 2009. Online: [http://www.acq.osd.mil/uas/docs/UMSIntegratedRoadmap2009.pdf]
4. P.W. Singer, "In the Loop? Armed Robots and the Future of War," *Defense Industry Daily*, January 28, 2009. Online:[http://www.defenseindustrydaily.com/In-the-Loop-Armed-Robots-and-the-Future-of-War-05267/]
5. "Drone attacks in Pakistan," *Wikipedia*, Online: [http://en.wikipedia.org/wiki/Drone_attacks_in_Pakistan]
6. Erik Sofge, "Can Robots Be Trusted?", *Popular Mechanics,* February 2010, pp. 54-61.
7. "Autonomy," *American Heritage Dictionary*, Online: [http://www.answers.com/topic/autonomy]
8. "Autonomous Agent," *Wikipedia*, Online: [http://en.wikipedia.org/wiki/Autonomous_agent]
9. "Autonomous Robot," *Wikipedia*, Online: [http://en.wikipedia.org/wiki/Autonomous_robot]
10. Diego Gambetta, ed., *Trust: Making and Breaking Cooperative Relations,* New York, NY: Basil Blackwell Ltd., 1988. Online: [http://www.nuffield.ox.ac.uk/users/gambetta/gambetta_trust%20book.pdf]
11. Cristiano Castelfranchi and Rino Falcone, "Principles of Trust for MAS: Cognitive Autonomy, Social Importance, and Quantification," Proceedings of ICMAS 1999. Online: [http://aois.org/99/castelfranchi-ICMAS-paper.rtf]
12. Sarvapali Ramchurn, "Devising a Trust Model for Multi-Agent Interactions using Confidence and Reputation," *Applied Artificial Intelligence*, 18:833-852, 2004. Online: [http://users.ecs.soton.ac.uk/nrj/download-files/jaai04.pdf]
13. Karen Fullam and K.Suzanne Barber, "Learning Trust Strategies in Reputation Exchange Networks," *The Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2006),* Hakodate, Japan, May 8-12, 2006, pp. 1241-1248. Online: [http://megatron.iiia.csic.es/art-testbed/pdf/FullamAAMAS2006.pdf]
14. *ART Testbed.* Online: [http://www.art-testbed.net/]
15. RoboCup Rescue, *Rescue Agents.* Online: [http://www.robocuprescue.org/agentsim.html]
16. ALADDIN, *Technologies.* Online: [http://www.aladdinproject.org/technologies.html]
17. Diplomacy AI Centre. Online: [http://www.daide.org.uk]
18. "Diplomacy (Game)," *Wikipedia,* Online: [http://en.wikipedia.org/wiki/Diplomacy_(game)]

Acronym List

AI	Artificial Intelligence
ALCM	Air Launched Cruise Missile
ART	Agent Reputation and Trust
DAIDE	*Diplomacy* AI Development Environment
DARPA	Defense Advanced Research Projects Agency
ISS	International Space Station
KPP	Key Performance Parameter
NASA	National Aeronautics and Space Administration
OSD	Office of the Secretary of Defense
SMXS	Software Maintenance Squadron
UAS	Unmanned Aerial System
UAV	Unmanned Aerial Vehicle
UGV	Unmanned Ground Vehicle
UMS	Unmanned Maritime System